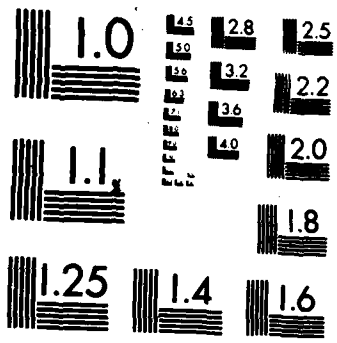
AD-A137 741 AN APPLICATION OF A MULTIDIMENSIONAL EXTENSION OF THE TWO-PARAMETER LOGIS..(U) AMERICAN COLL TESTING PROGRAM IOWA CITY IA R L MCKINLEY ET AL. AUG 83 RR-83-3-ONR 1/1

UNCLASSIFIED N00014-81-K-0817 F/G 12/1 NL

END
FILMED
3-84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

An Application of a Multidimensional Extension of the Two-Parameter Logistic Latent Trait Model

Robert L. McKinley
and
Mark D. Reckase

Research Report ONR83-3
August 1983

The American College Testing Program
Resident Programs Department
Iowa City, Iowa 52243

Prepared under Contract No. N00014-81-K0817
with the Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ONR83-3	2. GOVT ACCESSION NO. [illegible]	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Application of a Multidimensional Extension of the Two-Parameter Logistic Latent Trait Model		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Robert L. McKinley and Mark D. Reckase		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K0817
9. PERFORMING ORGANIZATION NAME AND ADDRESS The American College Testing Program P.O. Box 168 Iowa City, IA 52243		10. PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS P.E.: 61153N Proj.: RR042-04 T.A.: 042-04-01 W.V.: NR150-474
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research Arlington, VA 22217		12. REPORT DATE August 1983 13. NUMBER OF PAGES 40
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited. Reproduction in whole or part is permitted for any purpose of the United States government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Latent trait models
Multidimensional latent trait models
Multidimensional two-parameter logistic model

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A study was conducted to investigate the feasibility of a multidimensional IRT model. A two-parameter logistic IRT model and a multidimensional extension of that model were selected for this research. The design of the study employed two stages. The first stage consisted of generating simulation data to fit the multidimensional

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

model, applying the model to the data, and comparing the resulting estimates to the known parameters. The unidimensional model was also applied to these data. In addition to comparing the parameter estimates to the true parameters, the fit of the unidimensional and multidimensional models to the data were compared. The second stage of the study employed real response data. Items were selected from various subtests of a large test so as to simulate shorter tests with varying numbers of dimensions. Both models were applied to these data, and the resulting estimates were used to evaluate the fit of the models to the data. The results of the analyses of the simulation data indicated that the parameters of the multidimensional model could be accurately estimated. The results of the goodness of fit analyses indicated that the multidimensional model could more adequately model simulated multidimensional response data than could the unidimensional model. The results of the analyses of the real data indicated that the multidimensional model also more adequately modeled multidimensional real data than did the unidimensional model. It was concluded that the use of a multidimensional model does [redacted] seem to be feasible, and that more research was warranted.

CONTENTS

An Application of a Multidimensional Extension

of the Two-Parameter Logistic Latent Trait Model

Latent trait theory has become an increasingly popular area for research and application in recent years. Areas where latent trait theory has been applied include test scoring (Woodcock, 1974), criterion-referenced measurement (Hambleton, Swaminathan, Cook, Eignor, and Gifford, 1978), test equating (Marco, 1977; Rentz and Bashaw, 1977), tailored testing (McKinley and Reckase, 1980), and mastery testing (Patience and Reckase, 1978). While many of these applications have been successful, one unsolved problem is repeatedly encountered. Most latent trait models assume a unidimensional latent trait. As a result, applications of these models have been limited to areas in which the tests used measure predominantly one factor. When the assumption of unidimensionality is not met, such as is often the case with achievement tests, most latent trait models are inappropriate.

The purposes of the research presented here are to describe a latent trait model that is appropriate for use with tests that measure more than one dimension and to demonstrate its application to both real and simulated test data. In addition, procedures for estimating the parameters of the model will be presented.

The objectives of this research are to determine whether the proposed model more adequately explains multidimensional test data than does the undimensional version of the model, and to determine whether the results yielded by the application of the model are consistent with the results of another, more established multivariate data reduction procedure, factor analysis.

Method

The Model

The unidimensional model selected for this study, the two-parameter logistic (2PL) model, is given by

$$P_i(\theta_j) = \frac{\exp(Da_i(\theta_j - b_i))}{1+\exp(Da_i(\theta_j - b_i))} , \qquad (1)$$

where a_i is the discrimination parameter for item i, b_i is the difficulty parameter for item i, θ_j is the ability parameter for examinee j, and D=1.7.

The multidimensional model selected for this study, a multidimensional extension of the two-parameter logistic (M2PL) model, is given by

$$P_i(\theta_j) = \frac{\exp(d_i + a_i \theta_j)}{1+\exp(d_i + a_i \theta_j)} , \qquad (2)$$

where $P_i(\theta_j)$ is the probability of a correct response to item i by examinee j, d_i is a parameter related to the difficulty of item i, a_i is a vector of item discrimination parameters for item i, θ_j is a vector of ability parameters for examinee j, and

$$a_i \theta_j = \sum_{k=1}^{m} a_{ik} \theta_{jk} , \qquad (3)$$

where a_{ik} is the discrimination parameter for item i on dimension k, θ_{jk} is the ability parameter for examinee j on dimension k, and m is the number of dimensions modeled.

Estimation Procedures

The procedure used for item parameter estimation for the M2PL model is a modification of the marginal maximum likelihood procedure proposed by Bock and Aitkin (1981). Their procedure was modified to make it appropriate for use with the logistic distribution rather than with the normal distribution. The ability estimation procedure used for the M2PL model is a conditional maximum likelihood estimation procedure. It employs an iterative estimation routine based on the Newton-Raphson technique. A complete description of the ability estimation procedure is included in McKinley and Reckase (1983).

For the 2PL model, parameter estimation was performed using the LOGIST estimation program (Wood, Wingersky, and Lord, 1976). This procedure is the most commonly used procedure for estimating the parameters of the three-parameter logistic (3PL) model. It can be used for estimating the parameters of the 2PL model by holding the 'pseudo-guessing' parameter constant at zero.

Design

The general design of this study involved two stages. The first stage employed simulation data with known true item and person parameters. The second stage involved the use of real test data, sampled to have specified numbers of subtests in order to control to some degree the factor structure of the tests.

In the first stage of the study response data with one, two, and three dimensions were generated using the M2PL model and known parameters. The parameters of the unidimensional and multidimensional forms of the model were estimated for these data, and the resulting sets of estimates were compared to the true parameters and to each other.

In the second stage of the study actual response data for a large test with several subtests were sampled in such a way as to simulate tests having one, two, and three subtests. Although the tests were simulated, the item responses were actual item responses from an administration of the large test. The parameters of the 2PL and M2PL models were estimated, and the resulting estimates were compared with each other.

Datasets

Six datasets were employed in this research, three containing simulated item responses and three containing real item responses. One simulation dataset was generated to have one dimension, a second was generated to have two dimensions, and a third dataset was generated to have three dimensions. The first real dataset was constructed so as to have only one content area, the second had two content areas, and the third had three content areas.

The true item parameters for the simulation datasets were selected in the following way. The d-parameters were selected from a table of the standard normal distribution. They were sampled to have a mean of approximately zero and a standard deviation of approximately .5. The a-parameters, or discrimination parameters, were selected so that each item would have a high discrimination on only one dimension, and a low discrimination on the other two dimensions. For the unidimensional data only the d-values and the a-values for the first dimension were used for data generation. For the two-dimensional data the one-dimensional data item parameters were used along with the a-values for the second dimension. The three-dimensional data were generated using

the two-dimensional data item parameters along with the a-values for the third dimension. All three simulation datasets included data for 50 items and 1000 examinees.

For the real datasets, item responses were sampled from Form 16 of the Texas Grammar, Spelling, and Punctuation (GSP) test (University of Texas, 1978). For the real dataset having one content area, response data for the spelling subtest of the GSP test were sampled for 1000 examinees and 30 items. For the two-subtest dataset, data were sampled for 1000 examinees for 15 items from the spelling subtest and 15 items from the grammar subtest of the GSP test. For the three-subtest dataset, response data were sampled for 1000 examinees for 10 items from the spelling subtest, 10 items from the grammar subtest, and 10 items from the punctuation subtest of the GSP test. The items that were selected were those items having the highest factor loadings on the first factor from a principal components analysis performed on the individual subtests. The principal components analyses were performed on phi coefficients.

Analyses

Simulation Data Analyses The first analysis performed on the simulation data was to compare the item and person parameter estimates obtained for both the 2PL and the M2PL models to the known true parameters. To facilitate these and subsequent analyses, the item parameter estimates for the 2PL model were put in the M2PL form by multiplying the a- and b-values together to obtain a d-value. Of course, some differences in scale between the two models were still expected, due to the presence of the D term in the 2PL model. The d-parameter estimates were compared to each other and to the true d-parameters using Pearson product moment correlations. The correlations obtained for the two models were compared using a *t* - test (using Fisher's r to z transformation). For the unidimensional data the a-parameter estimates were compared to each other and to the true a-parameters using the same procedure.

For the multidimensional data there were different numbers of a-parameter estimates for the unidimensional and multidimensional forms of the model. Therefore, there was no one-to-one correspondence between the two sets of estimates. Because of this, correlations between the two sets of estimates would not be meaningful for evaluating the quality of the estimates. However, such correlations might lead to a better understanding of the relationship between the two forms of the model. Therefore, the intercorrelation

matrices for the a-parameter estimates were computed for the multidimensional data.

Another analysis performed on the simulation data was the computation, for each model, of a mean absolute deviation (MAD) statistic. This statistic is given by

$$MAD_i = \sum_{j=1}^{n} |P_{ij} - x_{ij}| \quad , \qquad (4)$$

where P_{ij} is the probability of a correct response to item i by examinee j given the item parameter estimates obtained for the model of interest, x_{ij} is the observed response to item i by examinee j, MAD_i is the mean absolute deviation statistic for item i, and n is the number of examinees. This statistic, an indicant of the ability of the models to predict item responses, was computed for all items for both the 2PL and M2PL models, and the mean MAD statistics for the two models were compared for the simulation data using analysis of variance techniques. In addition, a principal components solution was obtained on phi coefficients computed for each dataset. A varimax rotated factor solution was also obtained and used to facilitate the interpretation of the results of the other analyses. The number of factors rotated was equal to the number of dimensions used to generate the data.

Real Data Analyses For the real data the true parameters were not known. Therefore, the first analysis performed on the real data was the computation of the MAD statistics. The MAD statistics for the two models were once again compared using analysis of variance techniques. A principal components analysis was also performed for each of the real datasets, and the varimax rotated factor solution was used to facilitate interpretation of the results. The number of factors rotated was equal to the number of subtests included in the data.

Results

Simulation Data Analyses

True Item Parameters The true item parameters that were used to generate all of the simulation data are shown in Table 1. The d-parameters that are shown were used for all three simulation datasets. The first a-parameter column contains the item discrimination parameters used to generate the one-dimensional data. The second a-parameter column

contains the item discrimination parameters that, along with the first set of item discrimination parameters, were used to generate the two-dimensional dataset. The third column of a-parameters were used with the first two sets to generate the three-dimensional dataset.

Table 1
True Item Parameters Used to Generate Simulated Item Response Data

Item	d	a_1	a_2	a_3
1	0.35	1.40	0.30	0.15
2	-0.25	0.30	1.30	0.15
3	-1.15	0.10	0.30	1.65
4	-0.55	1.50	0.20	0.25
5	-0.05	0.35	1.35	0.20
6	1.00	0.15	0.30	1.60
7	-0.40	1.55	0.10	0.25
8	-0.70	0.40	1.70	0.15
9	0.30	0.40	0.25	1.75
10	-0.50	1.65	0.20	0.30
11	-0.10	0.20	1.30	0.15
12	1.05	0.35	0.15	1.60
13	-0.50	1.60	0.20	0.15
14	1.75	0.35	1.45	0.25
15	-1.10	0.20	0.15	1.40
16	0.10	1.75	0.20	0.35
17	-0.20	0.40	1.70	0.25
18	0.55	0.20	0.20	1.55
19	0.40	1.50	0.35	0.35
20	0.25	0.40	1.45	0.25
21	0.65	0.10	0.45	1.50
22	0.10	1.50	0.15	0.25
23	-0.35	0.30	1.60	0.25
24	-0.15	0.30	0.10	1.55
25	0.30	1.35	0.15	0.20
26	0.30	0.35	1.70	0.20
27	-0.30	0.15	0.30	1.75
28	0.40	1.60	0.40	0.25
29	-0.40	0.35	1.70	0.25
30	-0.40	0.35	0.15	1.70

Table 1(Continued)
True Item Parameters Used to Generate Simulated Item Response Data

Item	d	a_1	a_2	a_3
31	1.60	1.45	0.55	0.20
32	-1.00	0.15	1.45	0.10
33	-0.50	0.40	0.25	1.50
34	0.05	1.75	0.30	0.10
35	0.45	0.30	1.60	0.20
36	-0.30	0.30	0.30	1.50
37	-0.90	1.45	0.20	0.00
38	0.40	0.20	1.40	0.30
39	0.25	0.25	0.20	1.50
40	0.15	1.55	0.35	0.40
41	0.35	0.30	1.50	0.15
42	-0.35	0.15	0.30	1.70
43	-0.20	1.35	0.35	0.40
44	0.10	0.25	1.45	0.40
45	0.15	0.15	0.15	1.65
46	-0.15	1.70	0.25	0.20
47	0.35	0.15	1.70	0.10
48	-0.30	0.15	0.40	1.60
49	0.20	1.65	0.15	0.10
50	-0.30	0.40	1.55	0.35
Mean	0.00	0.70	0.68	0.67
S.D.	0.59	0.62	0.62	0.65

Table 1 also shows the means and standard deviations of the true item parameters. As can be seen, all of the item parameters had similar means and standard deviations. Dimensions 2 and 3 had mean a-values that were slightly lower than the mean a-values for dimension 1, with the dimension 3 a-values having the lowest mean. Dimension 3 also had the highest a-value standard deviation.

Table 2 shows the intercorrelation matrix for the item parameters shown in Table 1. As can be seen, there is no correlation between the true a-parameters and the true d-parameters ($r=0.03$ for dimensions 1 and 2, $r=-0.03$ for dimension 3). The a-parameters for the different dimensions were moderately negatively correlated. The a-parameters had correlations of -0.45 for dimensions 1 and 2, -0.51 for dimensions 1 and 3, and -0.50 for dimensions 2 and 3. The negative correlations among the a-values are a reflection of

the fact that items were simulated so as to have high a-values on only one dimension.

Table 2
Intercorrelation Matrix for the True Item Parameters Used to Generate the Simulated Item Response Data

Parameter	d	a_1	a_2	a_3
d	1.00	0.03	0.03	-0.03
a_1		1.00	-0.45	-0.51
a_2			1.00	-0.50
a_3				1.00

<u>Factor Analyses</u> Table 3 summarizes the results of the factor analyses performed on the simulation datasets that were generated using the item parameters shown in Table 1. For the one-dimensional data the factor loadings that are shown are for the first principal component from a principal components analysis of phi coefficients. For the two- and three-dimensional data the loadings shown are from a varimax rotated principal components solution.

For the one-dimensional data the first two eigenvalues from the principal components analysis were 6.54 and 1.34. These data appear to at least approximate unidimensionality. For the two-dimensional data the first three eigenvalues were 8.07, 4.03, and 1.25. These data clearly do not approximate unidimensionality. For the three-dimensional data the first four eigenvalues were 9.12, 4.51, 3.81, and 1.03. Again, these data are clearly not unidimensional.

Table 3
Factor Loadings Obtained for the One-, Two-, and Three-Dimensional Simulated Item Response Data

Item	One Dimensional	Two Dimensional		Three Dimensional		
	I	I	II	I	II	III
1	0.54	0.60	0.07	0.56	0.07	0.13
2	0.20	0.09	0.57	0.13	0.04	0.52
3	0.07	0.06	0.19	0.05	0.58	0.07
4	0.56	0.54	0.10	0.56	0.08	0.09
5	0.20	0.14	0.53	0.09	0.08	0.56
6	0.07	0.11	0.13	0.05	0.59	0.05
7	0.55	0.60	0.01	0.55	0.08	-0.02
8	0.22	0.11	0.57	0.12	0.04	0.60
9	0.22	0.25	0.12	0.11	0.64	0.03
10	0.62	0.58	0.11	0.61	0.09	0.03
11	0.12	0.10	0.56	0.08	0.08	0.54
12	0.18	0.12	0.05	0.13	0.54	0.08
13	0.55	0.57	0.06	0.55	0.06	0.10
14	0.15	0.06	0.47	0.16	0.04	0.49
15	0.08	0.06	0.04	0.11	0.57	0.08
16	0.62	0.60	0.07	0.58	0.11	0.15
17	0.25	0.14	0.59	0.10	0.10	0.61
18	0.19	0.15	0.13	0.08	0.58	0.04
19	0.57	0.58	0.15	0.58	0.13	0.05
20	0.18	0.17	0.50	0.16	0.01	0.53
21	0.03	-0.01	0.24	0.01	0.56	0.16
22	0.58	0.56	0.07	0.59	0.13	0.04
23	0.19	0.04	0.66	0.12	0.10	0.55
24	0.19	0.18	0.06	0.09	0.62	0.05
25	0.53	0.56	0.07	0.52	-0.02	0.10
26	0.21	0.16	0.60	0.11	0.06	0.61
27	0.12	0.05	0.13	0.01	0.62	0.08
28	0.60	0.55	0.14	0.61	0.02	0.15
29	0.17	0.14	0.62	0.11	0.04	0.60
30	0.23	0.17	0.16	0.14	0.60	0.05
31	0.48	0.50	0.09	0.50	0.07	0.15
32	0.00	0.04	0.54	0.02	0.04	0.49
33	0.24	0.23	0.08	0.11	0.57	0.09
34	0.64	0.63	0.08	0.63	0.04	0.15
35	0.17	0.11	0.62	0.10	0.07	0.56
36	0.17	0.16	0.15	0.12	0.58	0.07
37	0.54	0.56	0.06	0.55	0.12	0.09
38	0.16	0.09	0.58	-0.01	0.09	0.54
39	0.11	0.05	0.06	0.08	0.57	0.05
40	0.56	0.54	0.06	0.60	0.13	0.12

Table 3(Continued)
Factor Loadings Obtained for the One-, Two-, and Three-Dimensional Simulated Item Response Data

Item	One Dimensional	Two Dimensional		Three Dimensional		
	I	I	II	I	II	III
41	0.16	0.14	0.55	0.09	0.10	0.55
42	0.05	0.11	0.17	0.01	0.65	0.10
43	0.62	0.57	0.13	0.55	0.16	0.11
44	0.16	0.09	0.52	0.07	0.12	0.56
45	0.07	-0.03	0.10	0.04	0.61	0.01
46	0.62	0.64	0.03	0.61	0.02	0.09
47	0.11	0.03	0.56	0.07	0.03	0.60
48	0.09	0.07	0.25	0.05	0.62	0.16
49	0.60	0.60	0.01	0.63	0.00	0.14
50	0.19	0.21	0.60	0.09	0.12	0.56

Note. For the two- and three-dimensional data the factor loadings shown are from a varimax rotation of the principal components solution.

The correlations between the true a-parameters and the factor loadings shown in Table 3 are reported in Table 4. As can be seen from Table 4, there is a strong relationship between the discrimination parameters of the M2PL model and the factor loadings from the factor analysis solutions. The correlation of the a-parameter for the first dimension and the one-factor solution factor loadings was 0.98. For the two-factor solution the correlation between the a-parameters and the factor loadings was 0.98 for both dimensions. For the three-factor solution the correlation between the a-parameters and the factor loadings was 0.99 for the first dimension, as was the correlation between the a-parameters for the second dimension and the factor loadings for the third factor. The correlation between the a-parameters for the third dimension and the factor loadings for the second factor was also 0.99. As can be seen, the second and third factors in the three-subtest solution were reversed in order from the true parameters. There is also a strong relationship between the dimensionality of the data as determined by the eigenvalues and the dimensionality of the parameter vectors.

These analyses provide strong evidence for the validity of the procedure used to generate multidimensional item response data. They also provide some evidence that the M2PL model actually can be used to model multidimensional data. It remains to be seen whether the model is appropriate for realistic data. The next issue that must be addressed is whether the parameters of the model can be accurately estimated. This issue was addressed by the simulation data analyses to be reported next.

Table 4
Correlations of True Discrimination Parameters with the Varimax Rotated Factor Loadings for the Simulated Item Response Data

True Parameter	Factor Loadings					
	One Factor	Two Factor		Three Factor		
	I	I	II	I	II	III
a_1	0.98	0.98	-0.54	0.99	-0.51	-0.42
a_2	-0.43	-0.49	0.98	-0.47	-0.49	0.99
a_3	-0.51	-0.46	-0.40	-0.50	0.99	-0.54

<u>One-Dimensional</u> <u>Data</u> Table 5 shows the item parameter estimates that were obtained for both models for the one-dimensional simulation data. The means and standard deviations of the item parameter estimates are also shown in Table 5. Note that for the one-dimensional data, parameters were estimated for only one dimension using the M2PL model. As can be seen from the table, the estimates for the unidimensional simulation data were quite similar for the two models, although the mean discrimination parameter estimates were somewhat higher for the M2PL model than for the 2PL model. The correlation of the d-parameter estimate with the true d-parameter was .99 for both models. The correlation of the a-parameter estimates with the true a-parameter was .98 for the 2PL model and .99 for the M2PL model. The correlation of the two sets of d-parameter estimates was .99, as was the correlation between the two sets of a-parameter estimates.

Table 5
Item Parameter Estimates for the 2PL and M2PL Models for the One-Dimensional Simulated Item Response Data

Item	2PL		M2PL	
	d	a	d	a
1	0.23	0.89	0.32	1.12
2	-0.17	0.23	-0.29	0.31
3	-0.70	0.08	-1.18	0.14
4	-0.23	0.94	-0.41	1.21
5	-0.09	0.23	-0.16	0.32
6	0.52	0.10	0.88	0.11
7	-0.27	0.96	-0.44	1.17
8	-0.40	0.27	-0.68	0.37
9	0.16	0.26	0.25	0.36
10	-0.32	1.29	-0.48	1.42
11	-0.08	0.13	-0.14	0.21
12	0.52	0.21	0.87	0.31
13	-0.28	0.92	-0.48	1.18
14	1.14	0.23	1.91	0.32
15	-0.63	0.08	-1.09	0.15
16	0.10	1.20	0.13	1.36
17	-0.17	0.29	-0.30	0.44
18	0.32	0.22	0.54	0.31
19	0.23	0.99	0.33	1.23
20	0.10	0.21	0.16	0.29
21	0.41	0.04	0.69	0.06
22	0.02	1.01	0.00	1.23
23	-0.21	0.22	-0.36	0.31
24	-0.12	0.20	-0.21	0.29
25	0.24	0.84	0.38	1.10
26	-0.14	0.21	-0.25	0.34
27	-0.16	0.16	-0.27	0.16
28	0.29	1.13	0.44	1.34
29	-0.22	0.20	-0.38	0.28
30	-0.26	0.27	-0.44	0.38
31	0.94	0.88	1.44	1.10
32	-0.60	0.01	-1.02	0.02
33	-0.29	0.29	-0.50	0.41
34	-0.01	1.32	-0.07	1.46
35	0.27	0.20	0.45	0.28
36	-0.19	0.20	-0.33	0.28
37	-0.47	0.91	-0.77	1.18
38	0.32	0.17	0.53	0.26
39	0.12	0.12	0.20	0.16
40	0.06	0.94	0.08	1.18

Table 5(Continued)
Item Parameter Estimates for the 2PL and M2PL Models for the One-Dimensional Simulated Item Response Data

Item	2PL		M2PL	
	d	a	d	a
41	0.22	0.16	0.37	0.26
42	-0.16	0.08	-0.28	0.09
43	-0.18	1.23	-0.28	1.40
44	0.03	0.18	0.05	0.24
45	0.07	0.08	0.12	0.11
46	-0.11	1.21	-0.22	1.38
47	0.19	0.11	0.32	0.20
48	-0.15	0.10	-0.26	0.13
49	0.15	1.07	0.22	1.33
50	-0.16	0.20	-0.27	0.30
Mean	0.00	0.47	-0.02	0.59
S.D.	0.35	0.43	0.58	0.50

The great similarity of the estimates obtained for the two models was expected, since in the unidimensional case the two models are essentially the same model. Any differences that were found between the two sets of estimates were probably the result of differences between the two estimation procedures that were used. As indicated by the correlations that were obtained, the differences found between the two sets of estimates were minimal, involving primarily a difference in scale. The variance of the estimates for the 2PL model was less than the variance of the estimates for the M2PL model. A rescaling of the estimates to place them on the same scale might have eliminated most of the differences found between the two sets of estimates.

Descriptive statistics for the ability estimate distributions for both models for the one-dimensional simulation data are shown in Table 6. As can be seen, the statistics for both models are quite similar to the statistics for the true abilities. The one exception is the standard deviation of the M2PL ability estimates, which was much higher than the standard deviation of the 2PL estimates and the true abilities. The correlation of the estimates of ability with the true abilities was .91 for the 2PL model, and .92 for the M2PL model. The difference between these

two correlations was not significant. The correlation of the two sets of ability estimates was .99.

Table 6
Descriptive Statistics for the True and Estimated Ability Distributions for the 2PL and M2PL Models for the One-Dimensional Simulated Item Response Data

Statistic	True	2PL	M2PL
Mean	0.01	-0.01	0.02
Median	0.03	0.01	0.06
S.D.	1.02	1.03	1.60
Skewness	-0.04	-0.16	-0.07
Kurtosis	-0.18	0.24	-0.19

Two-Dimensional Data Table 7 shows the item parameter estimates that were obtained for both models for the two-dimensional simulation data. Also shown are the item parameter estimate means and standard deviations. The 2PL and M2PL item parameter estimate means are very similar, but the M2PL standard deviations are higher (and closer to the true values) than the 2PL standard deviations.

The intercorrelation matrix for the true and estimated item parameters for the two-dimensional simulation data is shown in Table 8. The parameter estimates for the multidimensional version of the model were quite strongly correlated with the true parameters. The correlation for the true and estimated d-parameter for the M2PL model was .99. For both a-parameters the correlation was .98. For the 2PL model the d-parameter estimate had a correlation of .98 with the true d-parameter, which was not significantly different from the correlation for the M2PL model. The two sets of d-parameter estimates had a correlation of .99. The unidimensional a-parameter estimates had a correlation of .47 with the first set of true a-parameters and .53 with the second set of true a-parameters. The correlations between the unidimensional a-parameter estimates and the multidimensional a-parameter estimates was .44 for the first set of a-parameter estimates, and .52 for the second set.

Table 7
Item Parameter Estimates for the 2PL and M2PL Models for the Two-Dimensional Simulated Item Response Data

Item	2PL		M2PL		
	d	a	d	a_1	a_2
1	0.18	0.71	0.25	1.32	0.13
2	-0.12	0.65	-0.31	0.08	1.26
3	-0.62	0.20	-1.08	0.12	0.29
4	-0.19	0.68	-0.46	1.12	0.27
5	0.09	0.69	0.09	0.20	1.14
6	0.68	0.23	1.14	0.17	0.23
7	-0.07	0.60	-0.25	1.26	0.07
8	-0.35	0.74	-0.74	0.13	1.30
9	0.21	0.31	0.32	0.42	0.18
10	-0.14	0.74	-0.38	1.23	0.21
11	-0.05	0.69	-0.17	0.16	1.16
12	0.60	0.15	1.01	0.21	0.10
13	-0.20	0.66	-0.51	1.21	0.13
14	0.90	0.58	1.61	0.09	1.10
15	-0.55	0.07	-0.95	0.13	0.10
16	0.09	0.74	0.04	1.31	0.18
17	-0.07	0.82	-0.20	0.18	1.27
18	0.30	0.23	0.47	0.24	0.21
19	0.38	0.86	0.55	1.30	0.34
20	0.22	0.70	0.32	0.30	1.08
21	0.34	0.20	0.66	-0.04	0.37
22	0.05	0.64	-0.03	1.17	0.14
23	-0.08	0.77	-0.24	-0.01	1.49
24	-0.11	0.18	-0.22	0.28	0.11
25	0.21	0.66	0.28	1.19	0.16
26	-0.18	0.89	-0.41	0.34	1.36
27	-0.14	0.16	-0.25	0.10	0.17
28	0.36	0.78	0.51	1.20	0.32
29	-0.20	0.90	-0.45	0.26	1.42
30	-0.20	0.28	-0.35	0.29	0.25
31	0.91	0.69	1.57	1.28	0.13
32	-0.53	0.60	-1.09	-0.05	1.25
33	-0.28	0.25	-0.52	0.37	0.13
34	0.07	0.80	0.03	1.39	0.21
35	0.23	0.82	0.38	0.23	1.38
36	-0.20	0.24	-0.36	0.25	0.25
37	-0.38	0.65	-0.84	1.23	0.12
38	0.34	0.73	0.60	0.13	1.27
39	0.19	0.09	0.32	0.08	0.08
40	0.21	0.64	0.28	1.13	0.12

Table 7(Continued)
Item Parameter Estimates for the 2PL and M2PL Models for the Two-Dimensional Simulated Item Response Data

Item	2PL		M2PL		
	d	a	d	a_1	a_2
41	0.19	0.76	0.28	0.22	1.19
42	-0.21	0.22	-0.39	0.16	0.23
43	0.02	0.78	-0.10	1.20	0.28
44	0.09	0.58	0.10	0.10	1.05
45	0.06	0.04	0.10	-0.03	0.14
46	-0.01	0.72	-0.18	1.43	0.06
47	0.24	0.57	0.45	-0.04	1.20
48	-0.22	0.26	-0.38	0.10	0.40
49	0.16	0.61	0.18	1.27	0.03
50	-0.12	1.02	-0.29	0.41	1.38
Mean	-0.04	0.55	0.01	0.54	0.55
S.D.	0.33	0.26	0.58	0.53	0.52

Table 8
Intercorrelation Matrix for the True and Estimated Item Parameters for the Two-Dimensional Simulated Item Response Data

Variable		True			2PL		M2PL		
		d	a_1	a_2	d	a	d	a_1	a_2
True	d	1.00	0.03	0.03	0.98	0.01	0.99	0.04	-0.03
	a_1		1.00	-0.45	0.12	0.47	0.05	0.98	-0.48
	a_2			1.00	0.01	0.53	0.01	-0.50	0.98
2PL	d				1.00	0.07	0.99	0.12	-0.05
	a					1.00	0.01	0.44	0.52
M2PL	d						1.00	0.06	-0.05
	a_1							1.00	-0.53
	a_2								1.00

Table 9 shows the descriptive statistics for the ability estimate distributions obtained for the two models for the two-dimensional simulation data. The statistics for the M2PL estimates were quite similar to the true parameter statistics, except that once again the standard deviation of the M2PL estimates was inflated. The 2PL statistics were much like the statistics for both dimensions of the true parameters, except that the 2PL estimate distribution was significantly leptokurtic (standard error for N=1000 is 0.155, $z = 6.823$, $p < .01$). This is probably due to an increased nonconvergence rate. For examinees for whom an ability estimate could not be obtained, the estimate was set equal to -4.0 or 4.0.

Table 10 shows the intercorrelation matrix for the true and estimated abilities for the two-dimensional simulation data. The correlations between the true ability parameters and the multidimensional estimates were .91 for both dimensions. The unidimensional ability parameter estimates had a correlation of .68 with the first set of true ability parameters and .70 with the first set of estimates for the M2PL model. The correlation between the unidimensional estimates and the second set of true ability parameters was .67, while a correlation of .73 was obtained for the unidimensional estimates and the second set of ability parameter estimates for the multidimensional model.

Table 9
Descriptive Statistics for the True and Estimated Ability Distributions for the 2PL and M2PL Models for the Two-Dimensional Simulated Item Response Data

Statistic	2PL	True		M2PL	
		θ_1	θ_2	θ_1	θ_2
Mean	0.02	0.10	0.02	0.12	0.07
Median	-0.01	0.08	0.01	0.16	0.10
S.D.	1.06	1.02	1.02	1.68	1.71
Skewness	0.15	0.12	-0.09	-0.02	-0.04
Kurtosis	1.06	0.08	0.20	0.00	-0.16

Table 10
Intercorrelation Matrix for the True and Estimated Ability Parameters for the Two-Dimensional Simulated Item Response Data

Variable		2PL	True θ_1	True θ_2	M2PL θ_1	M2PL θ_2
2PL		1.00	0.68	0.67	0.70	0.73
True	θ_1		1.00	0.04	0.91	0.11
	θ_2			1.00	0.04	0.91
M2PL	θ_1				1.00	0.06
	θ_2					1.00

<u>Three-Dimensional</u> <u>Data</u> Table 11 shows the item parameter estimates that were obtained for both models for the three-dimensional simulation data. The item parameter estimate means and standard deviations are also shown. As can be seen, the M2PL estimates once again have much higher standard deviations than the 2PL estimates. The 2PL a-value standard deviation is extremely low. The M2PL a-value standard deviations are much closer to the true values than the 2PL value, although the 2PL a-value mean is closer to the true value of 0.70 than the M2PL a-value means. Table 12 shows the intercorrelation matrix for the true and estimated item parameters for these data. Once again, the estimates for the M2PL model had high correlations with the true parameters. The d-parameter estimate had a correlation of .99 with the true d-parameter. The correlation of the first a-parameter estimate with the true first a-parameter was .98, as was the case for the second a-parameter. For the third set of a-parameters the correlation was .99. For the unidimensional version of the model, the correlation between the d-parameter and the estimated d-parameter was .99. The two sets of d-parameter estimates had a correlation of .99. The correlations obtained between the unidimensional a-parameter estimates and the three sets of true a-parameters were .69,-.26, and -.27, respectively. The corresponding correlations between the unidimensional a-parameter estimates and the three sets of multidimensional a-parameter estimates were .73, -.20, and -.27.

Table 11
Item Parameter Estimates for the 2PL and M2PL Models
for the Three-Dimensional Simulated Item Response Data

Item	2PL		M2PL			
	d	a	d	a_1	a_2	a_3
1	0.17	0.70	0.28	1.17	0.30	0.15
2	-0.12	0.53	-0.24	0.25	0.99	0.14
3	-0.54	0.55	-1.21	0.17	0.17	1.23
4	-0.29	0.66	-0.63	1.12	0.22	0.17
5	-0.01	0.56	-0.08	0.15	1.07	0.19
6	0.56	0.54	1.37	0.17	0.33	1.52
7	-0.19	0.52	-0.41	1.08	-0.03	0.11
8	-0.26	0.63	-0.61	0.27	1.35	0.21
9	0.17	0.61	0.35	0.23	0.04	1.34
10	-0.26	0.67	-0.62	1.36	-0.04	0.20
11	-0.10	0.54	-0.22	0.13	1.04	0.22
12	0.60	0.62	1.18	0.31	0.07	1.18
13	-0.33	0.65	-0.68	1.08	0.17	0.13
14	0.93	0.66	1.96	0.58	1.31	0.05
15	-0.54	0.63	-1.16	0.32	0.10	1.24
16	0.16	0.82	0.36	1.27	0.28	0.26
17	-0.10	0.67	-0.35	0.21	1.34	0.26
18	0.39	0.55	0.84	0.06	0.15	1.26
19	0.28	0.70	0.51	1.23	0.08	0.30
20	0.07	0.55	0.12	0.37	1.00	0.09
21	0.32	0.56	0.67	0.03	0.33	1.19
22	0.05	0.70	0.03	1.22	0.05	0.33
23	-0.19	0.63	-0.36	0.27	1.12	0.27
24	0.03	0.58	0.05	0.16	0.04	1.26
25	0.14	0.50	0.20	0.96	0.20	-0.02
26	-0.21	0.66	-0.49	0.23	1.40	0.23
27	-0.14	0.52	-0.31	0.05	0.15	1.27
28	0.22	0.76	0.43	1.32	0.31	0.00
29	-0.14	0.60	-0.36	0.26	1.24	0.21
30	-0.18	0.63	-0.44	0.27	0.09	1.27
31	0.92	0.77	1.73	1.21	0.38	0.00
32	-0.46	0.40	-0.93	0.01	0.97	0.05
33	-0.22	0.62	-0.48	0.32	0.24	1.14
34	0.01	0.80	-0.02	1.42	0.30	0.06
35	0.22	0.58	0.46	0.20	1.19	0.21
36	-0.14	0.60	-0.38	0.28	0.14	1.21
37	-0.48	0.70	-1.04	1.23	0.19	0.31
38	0.20	0.46	0.36	-0.06	1.03	0.19
39	0.16	0.52	0.36	0.22	0.04	1.19
40	0.07	0.84	0.03	1.27	0.26	0.31

Table 11(Continued)
Item Parameter Estimates for the 2PL and M2PL Models for the Three-Dimensional Simulated Item Response Data

Item	2PL		M2PL			
	d	a	d	a_1	a_2	a_3
41	0.14	0.56	0.26	0.19	1.07	0.20
42	-0.14	0.58	-0.25	0.08	0.21	1.46
43	-0.12	0.77	-0.30	1.10	0.23	0.35
44	0.03	0.59	-0.01	0.13	1.12	0.36
45	0.09	0.46	0.17	0.07	-0.01	1.27
46	-0.11	0.65	-0.30	1.30	0.10	0.08
47	0.21	0.54	0.41	0.06	1.20	0.12
48	-0.08	0.65	-0.22	0.00	0.38	1.40
49	0.19	0.72	0.31	1.39	0.25	-0.01
50	-0.17	0.62	-0.32	0.12	1.15	0.29
Mean	-0.16	0.62	0.00	0.54	0.51	0.53
S.D.	0.31	0.09	0.66	0.51	0.48	0.53

Table 12
Intercorrelation Matrix for the True and Estimated Item Parameters for the Three-Dimensional Simulated Item Response Data

Var.		True				2PL		M2PL			
		d	a_1	a_2	a_3	d	a	d	a_1	a_2	a_3
True	d	1.00	0.03	0.03	-0.03	0.99	0.12	0.99	0.07	0.06	-0.07
	a_1		1.00	-0.45	-0.51	0.02	0.69	-0.01	0.98	-0.43	-0.54
	a_2			1.00	-0.50	0.00	-0.26	0.00	-0.47	0.98	-0.45
	a_3				1.00	-0.02	-0.27	0.04	-0.48	-0.52	0.99
2PL	d					1.00	0.13	0.99	0.06	0.03	-0.02
	a						1.00	0.09	0.73	-0.20	-0.27
M2PL	d							1.00	0.03	0.03	0.00
	a_1								1.00	-0.43	-0.52
	a_2									1.00	-0.47
	a_3										1.00

Table 13 shows the descriptive statistics for the ability estimate distributions for both models for the three-dimensional simulation data The M2PL statistics are similar to the true statistics, except that the M2PL standard deviations are higher. Also, the M2PL dimension 1 kurtosis is significant (standard error=0.155, $z = 2.860$, $p < .01$), while the true value is not significant. The 2PL kurtosis is also significant ($z = 5.706$, $p < .01$), as is the 2PL skewness (standard error is 0.077, $z = 4.699$, $p < .01$). Again, the skewness and kurtosis of the ability estimate distributions are probably a reflection of nonconvergence.

Table 14 shows the intercorrelation matrix for the true and estimated ability parameters for the three-dimensional simulation data. The correlations between the three sets of ability parameter estimates for the M2PL model and the three sets of true ability parameters were .91, .90, and .90. The correlations obtained between the unidimensional ability parameter estimates and the three sets of true ability parameters were .57, .49, and .45. The corresponding correlations for the multidimensional estimates and the unidimensional estimates were .59, .48, and .48.

Table 13
Descriptive Statistics for the True and Estimated Ability Distributions for the 2PL and M2PL Models for the Three-Dimensional Simulated Item Response Data

Statistic	2PL	True			M2PL		
		θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
Mean	0.03	-0.01	0.00	0.05	0.06	0.02	0.01
Median	-0.01	0.02	0.01	0.03	0.01	0.08	-0.03
S.D.	1.07	0.98	0.99	1.02	1.47	1.59	1.64
Skewness	0.36	-0.05	0.01	0.07	0.07	-0.07	0.08
Kurtosis	0.88	0.10	-0.10	0.02	0.44	0.19	-0.07

Table 14
Intercorrelation Matrix for the True and Estimated Ability Parameters for the Three-Dimensional Simulated Item Response Data

Variable		2PL	True			M2PL		
			θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
2PL		1.00	0.57	0.49	0.45	0.59	0.48	0.48
True	θ_1		1.00	0.05	-0.03	0.91	0.05	0.00
	θ_2			1.00	-0.02	0.06	0.90	-0.01
	θ_3				1.00	0.02	-0.01	0.90
M2PL	θ_1					1.00	0.01	0.01
	θ_2						1.00	-0.06
	θ_3							1.00

<u>Overall Performance on Simulation Data</u> The final analysis that was performed on the simulation data was an analysis of variance performed on the MAD statistics. Table 15 shows the mean MAD statistics that were computed for both models for the simulation data. The standard deviations for these statistics are also shown. The dimensionality of the data and the model used were independent variables, with model as a repeated measures factor. The analysis of variance performed on these data yielded the results shown in Table 16.

Table 15
Descriptive Statistics for MAD Statistics Obtained for the Simulation Datasets

No. of Dimensions	Statistic	2PL	M2PL
1	Mean	0.43	0.41
	S.D.	0.06	0.08
2	Mean	0.43	0.36
	S.D.	0.04	0.07
3	Mean	0.43	0.31
	S.D.	0.02	0.03

Table 16
Two-Way Analysis of Variance on Mean Absolute Differences with Dimensionality of Data and Model as Independent Measures with Repeated Measures over Model

Source	SS	df	MS	F	p
Dimensionality	0.136	2	0.068	13.390	0.000
Error	0.749	147	0.005		
Model	0.390	1	0.390	1223.040	0.000
Model x Dim.	0.098	2	0.049	154.220	0.000
Error	0.047	147	0.000		

As can be seen, all of the effects were found to be significant. The test for the significance of the dimensionality effect yielded an $F = 13.39$, $p < .01$. Analysis of the cell means indicates that the models yielded lower mean MAD statistics as the dimensionality of the data increased. The test for the significance of the dimensionality by model interaction yielded an $F = 154.22$, $p < .01$. A look at the cell means, reported at the bottom of Table 15, reveals that the mean MAD statistics decreased at a much faster rate for the M2PL model than for the 2PL model. As the dimensionality of the data increased, then, the advantage gained by use of the multidimensional model increased.

The test for the model effect yielded an $F = 1223.04$, $p < .01$, indicating that across the three sets of response data the M2PL model yielded significantly lower mean MAD statistics. Paired t - tests were performed on these data to compare the mean MAD statistics yielded by the two models for each level of dimensionality. These t - tests yielded a $t = 10.64$, $p < .01$ for the unidimensional data, $t = 14.36$, $p < .01$ for the two-dimensional data, and $t = 46.30$, $p < .01$ for the three-dimensional data. Regardless of the dimensionality of the data, the M2PL model fit the data better than the 2PL model.

Real Data Analyses

Factor Analyses The results of the principal components analysis of phi coefficients for the three real data datasets are summarized in Table 17. For the two- and three-subtest data the factor loadings shown are from a varimax rotation of the principal components solution. The first two eigenvalues from the principal components analysis

of the one-subtest data are 4.22 and 1.78. The first three eigenvalues from the principal components analysis of the two-subtest data are 3.78, 2.27, and 1.24. For the three-subtest data the first four eigenvalues are 3.84, 2.72, 1.64, and 1.29.

Table 17
Factor Loadings Obtained for the One-, Two-, and Three-Subtest Real Item Response Data

Item	One Factor	Two Factor		Three Factor		
	I	I	II	I	II	III
1	0.16	0.57	-0.03	0.56	0.03	-0.07
2	0.15	0.59	0.04	0.60	0.08	0.01
3	0.16	0.37	0.12	0.45	0.17	0.05
4	0.36	0.46	0.16	0.40	0.06	0.00
5	0.25	0.33	0.10	0.58	0.05	0.06
6	0.13	0.36	0.03	0.48	0.05	0.00
7	0.20	0.42	0.02	0.52	0.10	0.07
8	0.24	0.56	0.02	0.48	-0.02	-0.01
9	0.28	0.32	0.08	0.66	-0.02	0.06
10	0.25	0.48	0.01	0.68	-0.02	0.03
11	0.52	0.53	0.08	0.07	0.46	0.05
12	0.32	0.47	-0.06	0.01	0.36	0.07
13	0.29	0.26	0.22	-0.10	0.30	0.10
14	0.55	0.61	-0.00	0.00	0.34	-0.03
15	0.39	0.64	-0.03	-0.01	0.14	0.80
16	0.34	0.11	0.47	0.11	0.21	0.26
17	0.22	0.03	0.38	0.03	0.52	0.14
18	0.49	0.06	0.25	0.07	0.36	0.09
19	0.36	0.01	0.28	-0.03	0.50	0.05
20	0.37	0.09	0.37	0.07	0.35	0.05
21	0.40	0.02	0.35	0.19	0.43	0.00
22	0.51	0.00	0.43	0.06	0.42	0.11
23	0.35	0.12	0.34	0.07	0.28	0.05
24	0.33	0.06	0.61	-0.07	0.43	0.10
25	0.46	0.15	0.36	0.04	0.46	-0.04
26	0.50	0.09	0.24	-0.00	0.10	0.78
27	0.40	0.07	0.46	0.15	0.39	0.19
28	0.34	0.05	0.27	0.08	0.23	0.29
29	0.54	0.02	0.49	0.04	0.44	0.01
30	0.57	0.10	0.33	-0.01	0.08	0.83

Note. For the two- and three-subtest data loadings are from a varimax rotated principal components solution.

As can be seen from the results of the factor analyses, the one-subtest data do at least approximate unidimensionality, even though some of the items did appear to load on specific factors. The first principal component is not a particularly large one, but it does seem to be dominant, as reflected by the smallness of the second component. The two-subtest data do not approximate unidimensionality. Rather, they seem to have two main components. This is a reasonable reflection of the subtest structure of these data. The factor loadings shown in Table 17 for the two-subtest data give an accurate picture of the subtest structure of the data, with the first 15 items having higher loadings on the first factor and the last 15 items having higher loadings on the second factor. The first 15 items were taken from the spelling test, and the last 15 were taken from the grammar test.

The three-subtest data results are not as clear. The first ten items were from the spelling test, the second ten were from the grammar test, and the last ten were from the punctuation test. From the results of the factor analysis it can be seen that the spelling items loaded on the first factor, and all of the second ten items except Item 15 loaded on the second factor. However, the last ten items, which were the punctuation items, tended to load on the second factor with the grammar items. This tendency is reflected in the smallness of the third eigenvalue from the principal components analysis. Only items 15, 26 and 30 had high loadings on the third factor. Thus, while the construction of the one- and two-subtest tests was successful, less success was achieved in constructing a three-subtest test.

<u>One-Subtest</u> <u>Data</u> The item parameter estimates that were obtained for the one-subtest data for both the 2PL and the M2PL models are shown in Table 18, along with their means and standard deviations. The two sets of d-values had similar standard deviations, but the 2PL model mean d-value was somewhat higher. The M2PL a-values had a higher mean and standard deviation than the 2PL a-values. Table 19 shows the intercorrelation matrix for the estimated item parameters for these data. The correlation of the two sets of d-parameters estimates was .93, and the correlation of the two sets of a-parameter estimates was .92.

Table 18
Item Parameter Estimates for the 2PL and M2PL Models for the One-Subtest Real Item Response Data

Item	2PL		M2PL	
	d	a	d	a
1	1.71	0.38	2.14	0.68
2	0.12	0.26	-0.04	0.22
3	0.31	0.28	0.24	0.25
4	2.10	0.87	1.86	1.03
5	0.58	0.55	0.42	0.29
6	0.51	0.23	0.56	0.30
7	-0.10	0.37	-0.55	0.38
8	0.23	0.51	-0.03	0.24
9	2.09	0.69	2.16	0.91
10	0.69	0.54	0.49	0.44
11	3.49	1.40	2.71	1.76
12	1.03	0.56	0.77	0.79
13	0.08	0.47	-0.59	0.74
14	2.14	1.26	1.37	2.04
15	1.04	0.70	0.46	1.01
16	0.54	0.59	-0.12	0.90
17	0.02	0.30	-0.46	0.51
18	1.03	1.09	0.16	1.49
19	1.16	0.69	0.69	1.01
20	1.46	0.71	1.12	1.01
21	1.66	0.70	1.28	1.24
22	2.24	1.14	1.43	1.41
23	0.72	0.56	0.09	1.01
24	1.28	0.67	1.02	0.80
25	2.30	1.07	1.69	1.35
26	1.40	0.97	0.48	1.32
27	2.78	0.90	2.80	1.42
28	-0.08	0.66	-1.08	0.94
29	2.70	1.26	1.83	1.65
30	2.66	1.45	1.44	1.69
Mean	1.26	0.73	0.81	0.96
S.D.	0.98	0.34	0.99	0.50

Table 20 shows the descriptive statistics for the ability estimate distributions for both models for the one-subtest data. The two distributions appear to be quite similar. The two sets of ability estimates had a correlation of .86.

Table 19
Intercorrelation Matrix for the Estimated Item Parameters for the 2PL and M2PL Models for the One-Subtest Real Item Response Data

Variable		2PL d	2PL a	M2PL d	M2PL a
2PL	d	1.00	0.82	0.93	0.80
	a		1.00	0.57	0.92
M2PL	d			1.00	0.56
	a				1.00

Table 20
Descriptive Statistics for the Ability Estimate Distributions for the 2PL and M2PL Models for the One-Subtest Real Item Response Data

Statistic	2PL	M2PL
Mean	0.08	0.11
Median	-0.08	-0.13
S.D.	1.18	1.19
Skewness	0.82	1.06
Kurtosis	1.89	2.04

<u>Two-Subtest</u> <u>Data</u> Table 21 shows the item parameter estimates that were obtained for the two models for the two-subtest real data, along with their means and standard deviations. The two sets of d-values are similar, though the 2PL mean is slightly higher and its standard deviation a little lower. The 2PL a-value mean was similar to the mean for the dimension 1 a-values for the M2PL model, while the standard deviation was more like the standard deviation for dimension 2 of the M2PL model. Dimension 2 of the M2PL model had a lower mean and standard deviation than dimension 1.

Table 22 shows the intercorrelation matrix for the two sets of item parameter estimates for these data. The correlation of the two sets of d-parameter estimates was .96. The correlation of the unidimensional a-parameter estimates with the multidimensional a-parameter estimates was .87 for the first dimension and -.40 for the second

dimension.

Table 21
Item Parameter Estimates for the 2PL and M2PL Models for the Two-Subtest Real Item Response Data

Item	2PL		M2PL		
	d	a	d	a_1	a_2
1	3.49	1.42	3.17	1.42	0.12
2	2.05	1.16	1.45	1.36	0.20
3	0.99	0.59	0.66	0.79	0.28
4	0.89	0.83	0.22	0.93	0.20
5	1.08	0.54	0.97	0.66	0.17
6	1.36	0.56	1.40	0.78	0.16
7	1.60	0.63	1.59	0.97	0.21
8	2.17	1.08	1.80	1.26	0.17
9	0.67	0.42	0.43	0.65	0.17
10	2.13	0.91	2.05	1.00	0.16
11	1.37	0.93	0.80	1.11	0.14
12	2.78	0.92	3.22	1.13	-0.08
13	-0.07	0.48	-0.94	0.69	0.41
14	2.72	1.28	2.27	1.39	0.23
15	2.32	1.13	2.04	1.57	0.01
16	1.15	0.79	1.12	0.47	0.89
17	-1.20	0.50	-2.68	0.33	0.85
18	0.22	0.31	-0.06	0.20	0.41
19	1.09	0.32	1.44	0.02	0.57
20	0.32	0.28	0.33	-0.22	0.51
21	1.35	0.49	1.58	0.10	0.73
22	0.18	0.42	-0.30	0.09	0.78
23	0.76	0.52	0.50	0.31	0.55
24	-0.07	0.87	-0.85	0.26	1.28
25	-0.05	0.18	-0.30	-0.26	0.61
26	0.96	0.43	1.04	0.22	0.32
27	-0.02	0.53	-0.79	0.25	0.73
28	0.53	0.30	0.40	0.20	0.47
29	-0.65	0.60	-1.66	0.07	0.80
30	1.13	0.54	1.11	0.26	0.55
Mean	1.04	0.67	0.73	0.60	0.42
S.D.	1.06	0.32	1.32	0.52	0.31

Table 22
Intercorrelation Matrix for Estimated Item Parameters for the 2PL and M2PL Models for the Two-Subtest Real Item Response Data

Variable		2PL d	2PL a	M2PL d	M2PL a_1	M2PL a_2
2PL	d	1.00	0.74	0.96	0.77	-0.70
	a		1.00	0.55	0.87	-0.40
M2PL	d			1.00	0.61	-0.67
	a_1				1.00	-0.72
	a_2					1.00

Table 23 shows the ability estimate distribution descriptive statistics for both models for the two-subtest real data. The 2PL distribution is similar to the distribution of M2PL ability estimates on dimension 2, although it was less leptokurtic. The dimension 1 M2PL estimates had a greater standard deviation, were more skewed, and were less leptokurtic than the dimension 2 or 2PL estimates.

Table 24 shows the intercorrelation matrix for the estimated ability parameters for the two-subtest real data. The correlation of the 2PL ability parameter estimates with the M2PL ability parameter estimates was .53 for the first dimension and .67 for the second dimension.

Table 23
Descriptive Statistics for Ability Estimate Distributions for the 2PL and M2PL Models for the Two-Subtest Real Item Response Data

Statistic	2PL	M2PL θ_1	M2PL θ_2
Mean	0.05	0.40	0.08
Median	-0.08	0.10	0.02
S.D.	1.10	1.60	1.21
Skewness	0.58	0.80	0.50
Kurtosis	1.09	0.67	1.83

Table 24
Intercorrelation Matrix for the True and Estimated Ability Parameters for the Two-Subtest Real Item Response Data

Variable		2PL	M2PL θ_1	M2PL θ_2
2PL		1.00	0.53	0.67
M2PL	θ_1		1.00	-0.12
	θ_2			1.00

<u>Three-Subtest</u> <u>Data</u> Table 25 shows the unidimensional and multidimensional item parameter estimates that were obtained for the three-subtest real item response data, along with their means and standard deviations. The 2PL d-values had a higher mean and a lower standard deviation than the M2PL d-values. The 2PL a-values had a higher mean and a lower standard deviation than dimensions 1 and 3 of the M2PL model. The 2PL a-value standard deviation was about the same as the M2PL dimension 2 a-value standard deviation. Table 26 shows the intercorrelation matrix for the two sets of item parameter estimates for these data. The two sets of d-parameter estimates had a correlation of .99. The correlation between the unidimensional a-parameter estimates and the multidimensional a-parameter estimates was .70 for the first dimension , -.38 for the second dimension, and .04 for the third.

Table 27 shows the descriptive statistics for the ability estimate distributions for both models for the three-subtest real data. The 2PL distribution is similar to the dimension 2 distribution for the M2PL model, although the 2PL standard deviation is somewhat smaller. The dimension 1 and 3 M2PL distributions have much higher standard deviations and are less skewed and leptokurtic. In addition, the dimension 1 mean is much higher.

Table 25
Item Parameter Estimates for the 2PL and M2PL Models
for the Three-Subtest Real Item Response Data

Item	2PL d	2PL a	M2PL d	M2PL a_1	M2PL a_2	M2PL a_3
1	3.08	1.21	3.28	1.13	0.42	-0.09
2	1.78	0.88	1.60	.18	0.23	-0.05
3	0.78	0.59	0.62	0.99	0.32	0.01
4	1.48	0.48	1.71	0.80	0.30	0.07
5	2.02	0.95	1.87	1.11	0.25	0.11
6	1.99	0.79	2.16	0.87	0.21	0.02
7	1.22	0.71	0.92	0.96	0.16	0.09
8	2.61	0.81	3.27	1.03	0.09	-0.06
9	2.51	1.14	2.44	1.33	0.20	0.22
10	2.05	0.89	2.16	1.42	0.12	0.13
11	1.11	0.75	1.07	0.34	0.95	0.24
12	-1.21	0.41	-2.65	0.11	0.66	0.35
13	0.31	0.28	0.27	-0.21	0.38	0.29
14	1.32	0.46	1.60	0.06	0.63	0.08
15	0.19	0.76	-0.70	0.01	0.51	2.00
16	0.75	0.53	0.49	0.29	0.40	0.34
17	-0.10	0.69	-1.03	0.14	0.79	0.32
18	-0.05	0.40	-0.68	0.19	0.48	0.20
19	-0.67	0.53	-1.64	-0.13	0.74	0.16
20	1.13	0.57	1.06	0.19	0.64	0.10
21	0.14	0.59	-0.62	0.46	0.61	0.06
22	0.29	0.53	-0.08	0.21	0.80	0.27
23	1.63	0.63	1.78	0.26	0.56	0.22
24	0.01	0.41	-0.45	-0.15	0.60	0.18
25	1.38	0.68	1.38	0.10	0.83	0.02
26	0.40	0.70	-0.03	-0.02	0.33	1.66
27	0.73	0.77	0.42	0.44	0.69	0.33
28	0.67	0.53	0.41	0.14	0.34	0.42
29	0.86	0.56	0.61	0.17	0.79	0.17
30	0.63	0.80	0.46	-0.02	0.23	2.05
Mean	0.97	0.67	0.72	0.45	0.48	0.33
S.D.	0.98	0.21	1.37	0.49	0.24	0.55

Table 26
Intercorrelation Matrix for the Estimated Item Parameters for the 2PL and M2PL Models for the Three-Subtest Real Item Response Data

Variable		2PL		M2PL			
		d	a	d	a_1	a_2	a_3
2PL	d	1.00	0.73	0.99	0.76	-0.52	-0.32
	a		1.00	0.62	0.70	-0.38	0.04
M2PL	d			1.00	0.69	-0.50	-0.33
	a_1				1.00	-0.64	-0.42
	a_2					1.00	-0.05
	a_3						1.00

Table 28 shows the intercorrelation matrix for the estimated ability parameters for the three-subtest real data. The unidimensional ability parameter estimates had a correlation of .33 with the first dimension of the multidimensional ability parameter estimates, .53 with the second dimension, and .47 with the third.

Table 27
Descriptive Statistics for the Ability Estimate Distributions for the 2PL and M2PL Models for the Three-Subtest Real Response Data

Statistic	2PL	M2PL		
		θ_1	θ_2	θ_3
Mean	0.12	0.61	0.31	0.24
Median	-0.04	0.26	0.09	0.09
S.D.	1.18	1.79	1.48	1.98
Skewness	0.86	0.64	0.90	0.10
Kurtosis	1.12	-0.07	1.03	0.32

Table 28
Intercorrelation Matrix for the True and Estimated Ability Parameters for the Two-Subtest Real Item Response Data

Variable		2PL	M2PL θ_1	M2PL θ_2	M2PL θ_3
2PL		1.00	0.33	0.53	0.47
M2PL	θ_1		1.00	-0.11	0.11
	θ_2			1.00	-0.25
	θ_3				1.00

Overall Performance on Real Data Table 29 shows the means and standard deviations of the MAD statistics that were computed for the real data. Table 30 summarizes a two-way analysis of variance that was performed on the MAD statistics computed for the three sets of real data. The number of subtests and the model used were the independent variables. Model was treated as a repeated measures variable.

Table 29
Descriptive Statistics for MAD Statistics Obtained for the Real Datasets

No. of Subtests	Statistic	2PL	M2PL
1	Mean	0.30	0.34
	S.D.	0.14	0.12
2	Mean	0.31	0.29
	S.D.	0.12	0.11
3	Mean	0.31	0.28
	S.D.	0.12	0.11

Table 30
Two-Way Analysis of Variance on Mean Absolute Differences with Number of Subtests and Model as Independent Measures with Repeated Measures over Model

Source	SS	df	MS	F	p
No. of Subtests	0.036	2	0.018	0.670	0.516
Error	2.355	87	0.027		
Model	0.007	1	0.007	7.730	0.007
Model x Subtests	0.066	2	0.033	37.250	0.000
Error	0.077	87	0.001		

As can be seen in Table 30, the number of subtests effect was not significant. However, the model effect was significant($\underline{F} = 7.73$, $\underline{p} < .01$), as was the model by number of subtests interaction($\underline{F} = 37.25$, $\underline{p} < .01$). Paired $\underline{t}$ -tests performed for each level of subtest structure yielded a $\underline{t} = 5.10$, $\underline{p} < .01$ for the one-subtest data, $\underline{t} = 3.62$, $\underline{p} < .01$ for the two-subtest data, and $\underline{t} = 5.96$, $\underline{p} < .01$ for the three-subtest data. It can be seen from the cell means shown in Table 29 that the 2PL model yielded a lower mean MAD statistic for the one-subtest data, while the M2PL model yielded lower mean MAD statistics for the two- and three-subtest data. Over all datasets, the M2PL model outperformed the 2PL model, although the estimation procedure used for the 2PL model seemed to perform better on the real data than did the estimation procedure for the M2PL model, as was reflected in the results of the analyses of the one-subtest data. The advantage of using the M2PL model became evident when two-subtest data were analyzed, and the advantage increased as the number of subtests increased.

Discussion

The purpose of this study was to investigate the feasibility of a multidimensional latent trait model. Several research questions were of interest. First, it was necessary to determine whether the parameters of the M2PL model could be accurately estimated. No model is useful if the parameters of the model cannot be accurately estimated.

A second research question addressed by this study is whether a multidimensional latent trait model more adequately models multidimensional item response data than

does a unidimensional model. If it does not, then it is not useful even if the parameters of the model can be estimated.

This research was divided into two parts: one part based on simulation data, and one part based on real data. The simulation part of the research was designed to determine whether the M2PL model could be used to model multidimensional item response data, whether the model parameters could be successfully estimated, and whether the model would fit multidimensional simulation data more adequately than the unidimensional version of the model. The real data part of the study was designed to determine whether the M2PL model would yield satisfactory results when applied to real data. The results of the simulation part of the study will be discussed first, and then a discussion of the real data part of the study will be presented.

Simulation Data Analyses

Factor Analysis Results The results of the factor analyses of the simulation data indicated that the attempt to generate multidimensional item response data was successful. There was a clear correspondence between the number of dimensions of the model parameters used to generate the data and the dimensionality of the data as indicated by the factor analyses. In addition, there was a clear relationship between the item discrimination parameters and the factor loadings obtained from the principal components analysis of phi coefficients. Thus, not only was the generation of the data successful, but evidence was obtained for the validity of the M2PL model.

One-Dimensional Data In the one-dimensional case the 2PL and M2PL models were essentially the same model. The M2PL model was just a reformulation of the 2PL model. Therefore, any differences found between the two models in the unidimensional case are probably due to differences in the estimation procedures used for the two models.

Even if the two estimation procedures yielded equal quality estimates, some differences might appear in the mean absolute differences for the two models. The M2PL procedure tends to yield estimates having greater variance than the estimates yielded by the 2PL procedure. More extreme estimates tend to yield predicted probabilities of responses that are more extreme (closer to 0 or 1), thus reducing the deviations between the item responses and predicted probabilities. It is unclear at this point whether there are inherent advantages in using one estimation procedure or the other. Any differences that do occur due to differences

in the estimation procedures will be evident in the results of the analyses of the unidimensional data, since for this case the two models are the same. Any differences found between the two models for the unidimensional case will serve as a baseline for evaluating the results of the analyses of multidimensional data.

The results of the analyses of the one-dimensional simulation data indicate that the the M2PL model performed slightly better than the 2PL model. The correlations of the true and estimated parameters were not significantly different for the two models, but the analyses of the mean absolute differences computed for the two models indicated that the goodness of fit of the M2PL model to the data was significantly better than the fit for the 2PL model. Although the parameter estimates were quite similar for the two models, the M2PL model estimation procedure yielded better fit to the data than the unidimensional estimation procedure did. The differences in mean absolute differences for the one-dimensional data serves as a baseline for evaluating the results of the analyses of the two- and three-dimensional data. If there is any advantage to using a multidimensional model, the difference between the mean absolute differences for the two models must be greater for the two- and three-dimensional data than for the unidimensional data.

Two-Dimensional Data The results of the analyses of the two-dimensional simulation data indicate that there is some advantage to using the M2PL model. The correlations of the estimated and true parameters for the M2PL model indicate that for two-dimensional simulation data the parameters of the model can be accurately estimated. The mean absolute differences analyses indicate that the M2PL model yields significantly better goodness of fit to the two-dimensional data than the unidimensional model. It is unclear how much of the difference between the two models is due to differences in the estimation procedures, but the results of the analyses of the unidimensional data indicate that at least part of the difference is due to differences in the estimation procedures for the two models.

Three-Dimensional Data As was the case for the two-dimensional data, for the three-dimensional data the M2PL model yielded parameter estimates that were highly correlated with the true parameters. From these results it appears that even with higher dimensionality the parameters of the M2PL model can be accurately estimated. The mean absolute differences analyses indicate that the M2PL model yields better fit to the three-dimensional data than the 2PL

model. Again, at least part of the difference between the two models is due to differences in the estimation procedures.

Overall Performance on Simulation Data It is clear that using the M2PL model for the multidimensional simulation data yields much better fit of the model to the data than could be obtained using the unidimensional model. For the unidimensional case there is very little difference between the two models, but as the dimensionality of the data increases so do the advantages of using the M2PL model model. Of course, these conclusions are based on the analysis of simulation data generated to fit the M2PL model. Any final conclusions regarding the value of using the M2PL model must be based not only on the results of simulation data analyses, but also on the results of real data analyses.

Real Data Analyses

Factor Analysis Results The results of the factor analyses performed on the real data indicate that the attempt to construct realistic multidimensional data was successful. The one-subtest data had one dominant factor, and the two-subtest data had two roughly equal factors. The three-subtest data had two large factors and a third smaller factor. Thus, with the exception of the smallness of the third factor of the three-subtest data, the factor structure of the real data closely paralleled the subtest structure of the data.

One-Subtest Data For the one-subtest real data the fit of the 2PL model to the data was better than the fit of the M2PL model. The estimation procedure used for the 2PL model appears to be more robust to violations of the assumptions of the model that are found in real data than is the case for the estimation procedure used for the M2PL model.

Two-Subtest Data The results of the analyses of the two-subtest data indicate that the fit of the M2PL model to these data was significantly better than the fit of the 2PL model. Thus, the advantages of using a multidimensional model with multidimensional real data are sufficient to overcome any advantage the 2PL model may have had on the basis of the estimation procedures.

Three-Subtest Data The results of the analyses of the three-subtest data were consistent with the results of the two-subtest data analyses. The fit of the M2PL model to the three-subtest data was better than the fit of the 2PL model

to the data. This is consistent with the results of the simulation data analyses.

Overall Performance on Real Data The analyses of the one-subtest data indicate that the estimation procedure used for the 2PL model may be somewhat better than the procedure used for the M2PL model when applied to real data. However, whatever disadvantage the M2PL model may have had due to the estimation procedure was overcome when the models were applied to multidimensional data. As the number of subtests in the real data increased, the difference in the fit of the two models to the data also increased.

Summary and Conclusions

The primary objective of the present research was to investigate the feasibility of a multidimensional latent trait model. The motivation behind this research was a desire to determine whether the great benefits realized through the use of unidimensional latent trait models could also be realized with a multidimensional model. A two-parameter logistic latent trait model and its multidimensional extension were selected for this research.

The design of the study employed two stages. The first stage consisted of generating simulation data to fit the multidimensional extension of the two-parameter logistic (M2PL) model, applying the model to the data, and comparing the resulting estimates with the known parameters. The unidimensional two-parameter logistic (2PL) model was also applied to these data. In addition to comparing the estimated parameters with the true parameters, the fit of the 2PL and M2PL models to the data were compared. The second stage of the study employed real response data. Items were selected from various subtests of a larger test that had been administered to a large sample in such a way as to simulate shorter tests with varying numbers of subtests. The 2PL and M2PL models were applied to these data, and the resulting estimates were used to evaluate the fit of the models to the data. The fit of the two models to the data were then compared to determine whether the M2PL model more adequately modeled the real data than did the 2PL model.

The results of the analysis of the simulation data indicated that the parameters of the M2PL model could be accurately estimated. The results of the goodness of fit analyses indicated that the M2PL model could more adequately model simulated multidimensional response data than did the

2PL model. The increase in dimensionality of the simulation data did not greatly reduce the accuracy with which the parameters of the M2PL model could be estimated.

The results of the analysis of the real test data indicated that the M2PL model also more adequately modeled multidimensional real data than did the 2PL model. The use of a M2PL model latent trait model does seem to be feasible, and the advantages gained by using such models seem to be great enough to warrant further research into this area.

REFERENCES

Bock, R. D. and Aitkin, M. Marginal maximum likelihood estimation of item parameters: An application of an EM algorthm. Psychometrika, 1981, 46, 443-459.

Hambleton, R. K., Swaminathan, H., Cook, L. L., Eignor, D. R., and Gifford, J. A. Developments in latent trait theory, models, technical issues, and applications. Review of Educational Research, 1978, 48, 467-510.

Marco, G. L. Item characteristic curve solutions to three intractable testing problems. Journal of Educational Measurement, 1977, 14, 139-160.

McKinley, R. L. and Reckase, M. D. A successful application of latent trait theory to tailored achievement testing (Research Report 80-1). Columbia: University of Missouri, Department of Educational Psychology, February 1980.

McKinley, R. L. and Reckase, M. D. MAXLOG: A computer program for the estimation of the parameters of a multidimensional logistic model. Behavior Research Methods and Instrumentation, 1983, 15, 389-390.

Patience, W. M. and Reckase, M. D. Self-paced versus paced evaluation utilizing computerized tailored testing. Paper presented at the annual meeting of the National Council on Measurement in Education, Toronto, 1978.

Rentz, R. R. and Bashaw, W. L. The National Reference Scale for Reading: An application of the Rasch model. Journal of Educational Measurement, 1977, 14, 161-180.

University of Texas. Grammar, Spelling, and Punctuation Test, Austin, TX, 1978.

Wood, R. L., Wingersky, M. S., and Lord, F. M. LOGIST: A computer program for estimating examinee ability and item characteristic curve parameters (Research Memorandum RM-76-6). Princeton: Educational Testing Service, June 1976.

Woodcock, R. W. Woodcock Reading Mastery Test. Circle Pines, MN: American Guidance Service, 1974.

Navy

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Arthur Bachrach
Environmental Stress Program Center
Naval Medical Research Institute
Bethesda, MD 20014

1 Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152

1 Liaison Scientist
Office of Naval Research
  Branch Office, London
Box 39
FPO New York, NY 09510

1 Dr. Alexander Bory
Applied Psychology
Measurement Division
NAMRL
NAS Pensacola, FL 32508

1 Dr. Robert Breaux
NAVTRAEQUIPCEN
Code N-095R
Orlando, FL 32813

1 Dr. Robert Carroll
NAVOP 115
Washington , DC 20370

1 Chief of Naval Education and Training
Liason Office
Air Force Human Resource Laboratory
Operations Training Division
WILLIAMS AFB, AZ 85224

1 Dr. Stanley Collyer
Office of Naval Technology
800 N. Quincy Street
Arlington, VA 22217

1 CDR Mike Curran
Office of Naval Research
800 N. Quincy St.
Code 270
Arlington, VA 22217

1 Dr. Doug Davis
CNET
Pensacola, FL

Navy

1 Dr. Tom Duffy
Navy Personnel R&D Center
San Diego, CA 92152

1 Mike Durmeyer
Instructional Program Development
Building 90
NET-PDCD
Great Lakes NTC, IL 60088

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
Code P13
NPRDC
San Diego, CA 92152

1 Dr. Cathy Fernandes
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Jim Hollan
Code 14
Navy Personnel R & D Center
San Diego, CA 92152

1 Dr. Ed Hutchins
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. William L. Maloy (02)
Chief of Naval Education and Training
Naval Air Station
Pensacola, FL 32508

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr William Montague
NPRDC Code 13
San Diego, CA 92152

Navy

1 Bill Nordbrock
1032 Fairlawn Ave.
Libertyville, IL 60048

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

6 Personnel & Training Research Group
Code 442PT
Office of Naval Research
Arlington, VA 22217

1 Special Asst. for Education and
Training (OP-01E)
Rm. 2705 Arlington Annex
Washington, DC 20370

1 LT Frank C. Petho, MSC, USN (Ph.D)
CNET (N-432)
NAS
Pensacola, FL 32508

1 Dr. Bernard Rimland (01C)
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode, Director
Training Analysis & Evaluation Group
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Frederick Steinheiser
CNO - OP115
Navy Annex
Arlington, VA 20370

Navy

1 Mr. Brad Sympson
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Frank Vicino
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Edward Wegman
Office of Naval Research (Code 411S&P)
800 North Quincy Street
Arlington, VA 22217

1 Dr. Ronald Weitzman
Naval Postgraduate School
Department of Administrative
Sciences
Monterey, CA 93940

1 Dr. Douglas Wetzel
Code 12
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Mr John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
San Diego, CA 92152

Marine Corps

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Director, Office of Manpower Utilizatio
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

1 Major Frank Yohannan, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Mr. James Baker
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Kent Eaton
Army Research Institute
5001 Eisenhower Blvd.
Alexandria , VA 22333

1 Dr. Beatrice J. Farr
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Director, Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commander, U.S. Army Research Institute
for the Behavioral & Social Sciences
ATTN: PERI-BR (Dr. Judith Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 20333

1 Joseph Psotka, Ph.D.
ATTN: PERI-IC
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Army

1 Mr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joyce Shields
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Hilda Wing
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

1 Dr. Robert Wisher
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

1 Technical Documents Center
Air Force Human Resources Laboratory
WPAFB, OH 45433

1 U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Mr. Raymond E. Christal
AFHRL/MOE
Brooks AFB, TX 78235

1 Dr. Alfred R. Fregly
AFOSR/NL
Bolling AFB, DC 20332

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Dr. T. M. Longridge
AFHRL/OTE
Williams AFB, AZ 85224

1 Mr. Randolph Park
AFHRL/MOAN
Brooks AFB, TX 78235

1 Dr. Roger Pennell
Air Force Human Resources Laboratory
Lowry AFB, CO 80230

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

Air Force

1 3700 TCHTW/TTGHR
2Lt Tallarigo
Sheppard AFB, TX 76311

1 Lt. Col James E. Watson
HQ USAF/MPXOA
The Pentagon
Washington, DC 20330

1 Major John Welsh
AFMPC
Randolph AFB, TX

1 Dr. Joseph Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Department of Defense

12 Defense Technical Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC

1 Dr. Craig I. Fields
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

1 Jerry Lehnus
HQ MEPCOM
Attn: MEPCT-P
Fort Sheridan, IL 60037

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 Dr. Wayne Sellman
Office of the Assistant Secretary
of Defense (MRA & L)
2B269 The Pentagon
Washington, DC 20301

1 Major Jack Thorpe
DARPA
1400 Wilson Blvd.
Arlington, VA 22209

Civilian Agencies

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E Street NW
Washington, DC 20415

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Private Sector

1 Dr. James Algina
University of Florida
Gainesville, FL 326

1 Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

1 1 Psychological Research Unit
NBH-3-44 Attn: Librarian
Northbourne House
Turner ACT 2601
AUSTRALIA

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Tel Aviv, Ramat Aviv 69978
Israel

1 Dr. R. Darrell Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52243

1 Dr. Ernest R. Cadotte
307 Stokely
University of Tennessee
Knoxville, TN 37916

1 Dr. John B. Carroll
409 Elliott Rd.
Chapel Hill, NC 27514

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

Private Sector

1 Dr. Hans Crombag
Education Research Center
University of Leyden
Boerhaavelaan 2
2334 EN Leyden
The NETHERLANDS

1 Dr. Dattpradad Divgi
Syracuse University
Department of Psychology
Syracuse, NE 33210

1 Dr. Fritz Drasgow
Department of Psychology
University of Illinois
603 E. Daniel St.
Champaign, IL 61820

1 Dr. Susan Embretson
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
Lawrence, KS 66045

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Benjamin A. Fairbank, Jr.
McFann-Gray & Associates, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Univ. Prof. Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

Private Sector

1 Dr. Dexter Fletcher
WICAT Research Institute
1875 S. State St.
Orem, UT 22333

1 Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01002

1 Dr. Robert Glaser
Learning Research & Development Center
University of Pittsburgh
3939 O'Hara Street
PITTSBURGH, PA 15260

1 Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

1 Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, MA 01002

1 Dr. Delwyn Harnisch
University of Illinois
242b Education
Urbana, IL 61801

1 Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 90010

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
603 East Daniel Street
Champaign, IL 61820

1 Dr. Jack Hunter
2122 Coolidge St.
Lansing, MI 48906

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

Private Sector

1 Dr. Douglas H. Jones
Advanced Statistical Technologies
Corporation
10 Trafalgar Court

Lawrenceville, NJ 08148

1 Professor John A. Keats
Department of Psychology
The University of Newcastle
N.S.W. 2308
AUSTRALIA

1 Dr. William Koch
University of Texas-Austin
Measurement and Evaluation Center
Austin, TX 78703

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat 23
9712GC Groningen
Netherlands

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Mr. Phillip Livingston
Systems and Applied Sciences Corporatio
6811 Kenilworth Avenue
Riverdale, MD 20840

1 Dr. Robert Lockman
Center for Naval Analysis
200 North Beauregard St.
Alexandria, VA 22311

Private Sector

1 Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

1 Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

1 Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

1 Dr. Scott Maxwell
Department of Psychology
University of Notre Dame
Notre Dame, IN 46556

1 Dr. Samuel T. Mayo
Loyola University of Chicago
820 North Michigan Avenue
Chicago, IL 60611

1 Mr. Robert McKinley
American College Testing Programs
P.O. Box 168
Iowa City, IA 52243

1 Dr. Barbara Means
Human Resources Research Organization
300 North Washington
Alexandria, VA 22314

1 Dr. Robert Mislevy
711 Illinois Street
Geneva, IL 60134

1 Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Ave., Fourth Floor
Redondo Beach, CA 90277

1 Dr. W. Alan Nicewander
University of Oklahoma
Department of Psychology
Oklahoma City, OK 73069

1 Dr. Melvin R. Novick
356 Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

Private Sector

1 Dr. James Olson
WICAT, Inc.
1875 South State Street
Orem, UT 84057

1 Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Cirle, NW
Washington, DC 20036

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

1 Dr. Thomas Reynolds
University of Texas-Dallas
Marketing Department
P. O. Box 688
Richardson, TX 75080

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 Frank L. Schmidt
Department of Psychology
Bldg. GG
George Washington University
Washington, DC 20052

1 Dr. Walter Schneider
Psychology Department
603 E. Daniel
Champaign, IL 61820

Private Sector

1 Lowell Schoer
Psychological & Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Kazuo Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

1 Dr. William Sims
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 Dr. Peter Stoloff
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. William Stout
University of Illinois
Department of Mathematics
Urbana, IL 61801

Private Sector

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Lab
252 Engineering Research Laboratory
Urbana, IL 61801

1 Dr. Maurice Tatsuoka
220 Education Bldg
1310 S. Sixth St.
Champaign, IL 61820

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

1 Dr. J. Uhlaner
Uhlaner Consultants
4258 Bonavita Drive
Encino, CA 91436

1 Dr. V. R. R. Uppuluri
Union Carbide Corporation
Nuclear Division
P. O. Box Y
Oak Ridge, TN 37830

1 Dr. David Vale
Assessment Systems Corporation
2233 University Avenue
Suite 310
St. Paul, MN 55114

1 Dr. Howard Wainer
Division of Psychological Studies
Educational Testing Service
Princeton, NJ 08540

1 Dr. Michael T. Waller
Department of Educational Psychology
University of Wisconsin--Milwaukee
Milwaukee, WI 53201

Private Sector

1 Dr. Brian Waters
HumRRO
300 North Washington
Alexandria, VA 22314

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 Dr. Rand R. Wilcox
University of Southern California
Department of Psychology
Los Angeles, CA 90007

1 Wolfgang Wildgrube
Streitkraefteamt
Box 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. Bruce Williams
Department of Educational Psychology
University of Illinois
Urbana, IL 61801

1 Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

END

FILMED

3-84

DTIC